Aggregation of Experts Judgments for Climate Tipping Points

Marcello Basili[1] and Federico Crudu[1,2]

[1]Department of Economics and Statistics, University of Siena, Piazza San Francesco 7, Siena, Italy.
[2]Centre For North South Economic Research, Via S. Giorgio 12, Cagliari, Italy.

Contributing authors: marcello.basili@unisi.it; federico.crudu@unisi.it;

Abstract

This paper introduces a method for the evaluation of the occurrence of tipping points based on the combination of probability intervals from experts judgments elicited in face-to-face interviews. The computation of such conditional probabilities is based on the aggregation of imprecise probability judgments through the Steiner point. The probability of a tipping point can be updated by the standard Bayes rule to generate tipping point scenarios. Our results suggest that tipping events may happen with relatively large probabilities, in contrast with the view that tipping points are low-probability-high-impact events. This suggests that mitigation and containment policies cannot be further postponed.

Keywords: Bayesian updating; aggregation of opinions; global warming; judgmental forecasting; Steiner point; tipping points

JEL Classification: Q54; D81; C10

1 Introduction

As of today, there is a large literature suggesting that the consequences of climate change may involve abrupt changes and tipping points (e.g. Lenton et al., 2019).[1] Armstrong McKay et al. (2022) identify a finite core of tipping point elements able to modify the Earth system functioning and show that six climate tipping points global elements are likely even in the Paris Agreement range of 1.5°C to 2°C warming. The concept of tipping point was introduced into the scientific debate in the '80s of the last century by the IPCC to represent large scale discontinuities in the climate system. At that time, experts believed that tipping points would be crossed if global warming had exceeded 5°C. In the 2018 IPCC report (IPCC, 2018b) experts suggested that tipping points could be crossed even between 1 and 2°C of global warming.[2]

The main objective of this paper is to introduce a new methodology for the evaluation of occurrence of tipping points based on the combination of probability intervals from experts judgments elicited in face-to-face interviews. When experts face uncertainty or have imprecise knowledge about future states of the world, alternative approaches that differ from Bayesian pool methods are required. In this paper we apply an aggregation method based on the Steiner point introduced in Basili and Chateauneuf (2020). This method assumes that experts have imprecise information represented by convex sets of probability distributions with the requirement that the intersection of these sets is not empty and experts have at least one common probability distribution. A desirable feature of this approach is that the opinion pooling is the probability distribution of the tipping points we considered. This differs from other methods where some type of (simple or weighted) average is defined (see, e.g., Clemen & Winkler, 1999, 2007). Furthermore, by means of a standard Bayesian updating rule

[1] A tipping point is defined as "[a] level of change in system properties beyond which a system reorganizes, often abruptly, and does not return to the initial state even if the drivers of the change are abated. For the climate system, it refers to a critical threshold when global or regional climate changes from one stable state to another stable state" (IPCC, 2018b). See also the recent article by Armstrong McKay et al. (2022) for a review and a thorough definition of tipping point.

[2] Lenton et al. mention that some "models suggest that the Greenland Ice sheet could be doomed at 1.5°C of warming, which could happen as soon as 2030" (Lenton et al., 2019, p. 592).

we can provide posterior probabilities that may be used to generate future scenarios or hypotheses for future studies, once more data become available (see Section 3 for more details on the aggregation approach).[3] It is important to notice that the construction of the Steiner point requires that the intersection of the experts' opinion be non empty. However, if that were not the case, the aggregation could be applied to subsets of consistent experts (see Section 3.1).

To provide an assessment of the occurrence of tipping points we use the elicitation data in Kriegler et al. (2009).[4] In Kriegler et al.'s seminal paper, experts are asked to provide probability intervals about three temperature scenarios. In this context, intervals of probabilities are a representation of uncertain and imprecise judgements.[5] The aggregation method proposed in this paper is new in the context of problems related to climate change.

The paper proceeds as follows. Section 2 briefly discusses the related literature. Section 3 introduces the fundamental aspects of our theoretical framework. In Section 4 we present the conditional probabilities of occurrence of various tipping points computed via the Steiner point as well as the results of the Bayesian updating. Finally, Section 5 concludes the study.

2 Related Literature

Communicating uncertainty about climate change is crucial to effectively influence policy decisions and shape public opinion. In this context, the IPCC special report Global Warming of 1.5°C (IPCC, 2018a) is an important example. The IPCC special

[3]It seems that there are no follow up studies on experts' assessment of tipping points. This is also confirmed in the correspondence with leading researchers in the field. Other recent studies use the data and results in Kriegler, Hall, Held, Dawson, and Schellnhuber (2009), see, e.g., Gaucherel and Moron (2017); Wunderling, Donges, Kurths, and Winkelmann (2021).

[4]The data are collected from Kriegler et al. (2009) supplemental appendix and are available at this link along with the replication code.

[5]In Kriegler et al. (2009) the low temperature scenario (Low) considers an increase between about 1°C and 2°C by year 2200 in comparison with year 2000. The medium temperature scenario ($Medium$) considers an increase between about 2°C and 4°C by year 2200 in comparison with year 2000. The high temperature scenario ($High$) considers an increase between about 4°C and 8°C by year 2200 in comparison with year 2000.

report is based "on the assessment of around 6,000 peer-review publications, most of them published in the last few years" (IPCC, 2018a, p. v). The report aggregates multiple forms of knowledge to address and communicate the degree of certainty (or lack thereof) of specific findings.

In general, however, some critical issues emerge in the treatment of uncertainty: findings are based on multiple lines of evidence and are expressed using confidence qualifiers and many of them depend on certain model assumptions. In addition to that, findings have to be updated if new information becomes available. Borsuk and Tomassini (2005), Tomassini, Reichert, Knutti, Stocker, and Borsuk (2007), Knutti et al. (2008), Zickfeld et al. (2007), Zickfeld, Morgan, Frame, and Keith (2010), Kriegler et al. (2009) highlight that aggregation of probabilistic projections with a variety of statistical models is an unsolved problem. One fundamental challenge of the assessment process is to summarize such information into a unique quantity. However, due to the heterogeneity in sources and quality of the information, obtaining a unique synthetic measure may become a daunting task.

When formal statistical procedures are unavailable, expert judgment approaches are often employed to provide an assessment of uncertainty (see e.g. Mastrandrea et al., 2011).[6] Significantly, "one option is to resort to imprecise probability (e.g., Kriegler and Held (2005); Hall, Fu, and Lawry (2007); Tomassini et al. (2007)), that is, consider an uncertainty in PDFs [probability distribution functions] or sets of PDFs" (Knutti et al., 2008, p. 2658). If each expert has a set of probability distributions, the mean value for each scenario is evaluated along with all the individual PDFs (Tomassini et al., 2007).

Experts' quantitative judgments are often elicited in face-to-face interviews. Then, using a range of different procedures, mean (averaged over experts) ranks are computed (Kriegler et al., 2009; Zickfeld et al., 2010). In Zickfeld et al. (2010), 14 experts (leading

[6] In 1975, the U.S. Nuclear Regulatory Commission (NCR) introduced for the first time a procedure for the elicitation process and, since then, techniques and methods have spread to other areas such as volcanology, public health, ecology, aeronautics, climatology etc. (Cooke, 2013).

climate scientists) discuss about three scenarios (high, medium, low) of net radiative forcing at the top of the atmosphere from anthropogenic sources through the year 2200. Experts use a cardinal scale from 0 (no chance) to 1 (definite chance) for each of the three forcing trajectories. Experts elicit probabilities for each scenario and in the following step they are asked to estimate the median of the mean trajectories of warming between 2000 and 2050. What they find is that it falls between 0.16°C/decade and 0.36°C/decade.

Providing a realistic assessment of the probability of a tipping event may have a fundamental impact in influencing policy decisions concerning climate change. Unfortunately, there are very few models able to simulate abrupt climate changes as consequences of realistic external forcing and, even when sophisticated climate models show effects, in particular a temperature response, they are weak with respect to Dansgaard-Oeschger events.[7] To manage the complexity and the uncertainty implied in climate models, policy makers (PMs) have no alternative but to resort to experts and obtain experts' elicitations; that is, expert probabilistic judgments (Colson & Cooke, 2018), that are represented by imprecise probabilities or intervals of probabilities, generally. Once the elicitation is completed, the data are aggregated. Using a reliable aggregation method to provide a synthesis of the various opinions seems of paramount importance, also given the fact that no clear dominant approach exists.[8]

[7]Dansgaard-Oeschger events are rapid climate fluctuations such as the Greenland ice melting occurred in the Eemian interglacial. This event took the form of rapid warming episodes, followed by gradual cooling periods, that increased the average annual temperature on the Greenland ice sheet of 8 °C over 40 years (Dansgaard et al., 1993; Heinrich, 1988).

[8]See Lam and Majszak (2022) for a recent review on the challenges and opportunities of expert judgment for the assessment of climate tipping points.

3 Theoretical Framework

The aggregation process of probabilistic opinions (opinion pooling) entails a function, known as a pooling rule, that elicits a consensus distribution. Such a consensus distribution is determined among not necessarily independent and fully competent experts when each of them has multiple priors on future states of the world.

Under uncertainty or deep uncertainty experts have partial, incomplete or fuzzy knowledge and their beliefs cannot be represented by a unique, additive and fully reliable probability distribution, but either by a finite set of them, an interval of probabilities or by a non necessarily additive measure (e.g. a capacity) (e.g. Basili & Chateauneuf, 2011, 2016; Basili & Pratelli, 2015, and references therein).[9]

In this paper we assume that opinions are expressed through different probability distributions and that there exists a PM who adopts a multiple priors decision model. The set of probability distributions of all experts can be considered a reflection of the PM's assessment of the reliability of available information about the underlying uncertainty, that is, her perception of uncertainty; the elicited aggregation rule incorporates the PM's attitude about scanty and vague information. Facing the set of all probability distributions attached by experts to possible events, the PM evaluates their probability intersection. Such a common opinion is the Steiner point of the convex capacity that emerges from the aggregation of experts' opinions.

The aggregation via the Steiner point provides a natural and tractable way to summarize opinions and only assumes that experts have imprecise information but nothing about their competence, experience and independence. This is, for example, the case of non Bayesian groups, i.e. groups where experts do not have a unique and

[9]Deep uncertainty includes "situations in which we are still able (or assume) to bound the future around many possible plausible futures and situations in which we only know that we do not know" (Marchau, Walker, Bloemen, & Popper, 2019). See also Rohmer, Le Cozannet, and Manceau (2019) and Frederikse et al. (2020) for some recent discussion of the concepts of uncertainty in the context of climate change issues.

additive probability distribution.[10] In addition, the Steiner point not only allows us to derive the consensus distribution about a given event, but it is amenable to Bayesian updating. This means that the PM can simply update the elicited distribution when new information is made available without calling for a new round of interviews. Whenever experts have opinions that are too heterogeneous the aggregation cannot be performed (see the G-consistency condition below in this section). Yet, the Steiner point can be applied to subsets of individuals allowing us to identify different trends and common opinions among consistent experts. In general, the aggregation through the Steiner point suggests that the experts are consistent and that the interpretation of a tipping point as a high-impact-low-probability event seems to be incorrect (see also Lenton & Ciscar, 2013; Lenton et al., 2019). We find, in fact, that the probability of crossing a tipping point threshold is non trivially larger than zero in nearly all cases, even for low climate change scenarios. Furthermore, it has been claimed (Lenton & Ciscar, 2013) that a more adequate picture for the representation of tipping points would be to provide the (joint) probability distribution for tipping each element.

3.1 Experts' Consistency

Let us consider a finite set $\mathcal{S} = \{s_1, \ldots, s_n\}$ of states of the world and let $\Sigma = 2^{\mathcal{S}}$ be the σ-algebra associated to the set $\mathcal{S}$. Since the study deals with discrete events, P is a probability mass function (pmf) on $(\mathcal{S}, \Sigma)$ such that $P : \Sigma \to [0, 1]$. For a specific pmf, $p(s_i) = P(S = s_i)$ for any $i \in \{1, \ldots, n\}$ and $\sum_{i=1}^{n} p(s_i) = 1$.

Consider now a non-negative set function $v : \Sigma \to \mathbb{R}$. Such a function is a capacity (a non additive probability) on $(\mathcal{S}, \Sigma)$ if, for any $\mathcal{A}, \mathcal{B} \in \Sigma$, $\mathcal{A} \subseteq \mathcal{B} \Rightarrow v(\mathcal{A}) \leq v(\mathcal{B})$, $v(\emptyset) = 0$ and $v(\mathcal{S}) = 1$, where $\emptyset$ is the empty set. The capacity $v(\cdot)$ is convex if

[10] The experts that provide judgements for the IPCC assessment may be interpreted as a non Bayesian group, as they typically provide evaluations in terms of qualitative levels of confidence and quantitative likelihoods of occurrence of a given event generally expressed as intervals of probabilities. Mastrandrea et al. (2011) characterize the degree of certainty in key findings with two metrics: "[c]onfidence in the validity of a finding, based on the type, amount, quality, and consistency of evidence (e.g., mechanistic understanding, theory, data, models, expert judgment) and the degree of agreement, [c]onfidence is expressed qualitatively" and "[q]uantified measures of uncertainty in a finding expressed probabilistically (based on statistical analysis of observations or model results, or expert judgment)".

$v(\mathcal{A} \cup \mathcal{B}) + v(\mathcal{A} \cap \mathcal{B}) \geq v(\mathcal{A}) + v(\mathcal{B})$. The corresponding dual capacity is defined as $\bar{v}(\mathcal{A}) = 1 - v(\mathcal{A}^c)$. The dual capacity $\bar{v}(\mathcal{A})$ is concave. Uncertainty is modeled via the core of the convex capacity v. The core $\mathcal{C}(v)$ is a set of probability distributions P on $(\mathcal{S}, \Sigma)$ such that $P(\mathcal{A}) \geq v(\mathcal{A})$ $\forall \mathcal{A} \in \Sigma$.

The PM's problem is to learn the distribution of a certain event by aggregating experts' opinions. Let P_0 denote the unobserved pmf that governs the phenomenon under study. The PM asks expert j, $(j = 1, \ldots, m)$, to provide a lower and upper bound for the probability $p_i = p_0(s_i) = P_0(S = s_i)$. The set of possible probabilities considered by expert j is defined as

$$\mathcal{P}^j = \left\{ P^j = \left(p_1^j, \ldots, p_i^j, \ldots, p_n^j \right) : a_i^j \leq p_i^j \leq b_i^j, i = 1, \ldots, n \right\}.$$

The PM will accept the expert's opinion if two consistency conditions are met. Specifically, for the expert's opinions to make sense the bounds a_i^j and b_i^j must meet the following conditions of *individual consistency* (*I-consistency*).

Condition 1. [**I-consistency**] The bounds a_i^j and b_i^j satisfy the conditions $0 \leq a_i^j \leq b_i^j \leq 1$ and $\sum_{i=1}^n a_i^j \leq 1 \leq \sum_{i=1}^n b_i^j$.

The set $\mathcal{P}^j$ is not empty if and only if the I-consistency condition is met. Furthermore, $\mathcal{P}^j$ can be seen as the core $\mathcal{C}(v^j)$ of a given convex capacity v^j where

$$v^j(\mathcal{A}) = \max \left(\sum_{i \in \{i : s_i \in \mathcal{A}\}} a_i^j,\ 1 - \sum_{i \in \{i : s_i \notin \mathcal{A}\}} b_i^j \right) \tag{1}$$

and the set $\mathcal{A} \subseteq \mathcal{S}$ is a set of states of the world (Chateauneuf & Cornet, 2018; De Campos, Huete, & Moral, 1994). It is expected that the unknown distribution P_0 be in the the set $\mathcal{P}^j$ of the generic expert j, i.e. $P_0 \in \mathcal{P}^j$. In addition to that, it

is expected also that P_0 be in the intersection of all $\mathcal{P}^j$. Hence, the following *group consistency* (*G-consistency*) condition is supposed to be met.

Condition 2. [**G-consistency**] The intersection of the probability sets associated with the pool of experts is non empty, i.e., $\mathcal{P} = \cap_{j=1}^{m} \mathcal{P}^j \neq \emptyset$ and that $P_0 \in \mathcal{P}$.

Whenever the intersection set $\mathcal{P} = \emptyset$, this is, when experts have conflicting opinions, the PM may still be able to extract valuable information by applying the G-consistency condition to a subset of experts. This approach may reveal, for example, whether some experts are more or less optimistic about a given phenomenon.

3.2 Aggregation and Scenarios

The consensus distribution set is associated to a convex capacity $v(\cdot)$, defined as

$$v(\mathcal{A}) = \max\left(\sum_{i\in\{i:s_i\in\mathcal{A}\}} a_i,\ 1 - \sum_{i\in\{i:s_i\notin\mathcal{A}\}} b_i\right) \tag{2}$$

where $a_i = \max_j a_i^j$ and $b_i = \min_j b_i^j$. In this context the Steiner point is particularly relevant, since the Steiner point of $\mathcal{C}(v)$ is the center of the core of the convex capacity v and represents the consensus probability for the given set of experts. The Steiner point is denoted as $\Pi^v \in \mathcal{C}(v)$.

It is interesting to notice that when the set of states of the world is finite, the Steiner point coincides with the Shapley value (Basili & Chateauneuf, 2020; Pechersky, 2015; Shapley, 1971) and it is easily computed via the following expression

$$\Pi_i^v = \sum_{s_i\in\mathcal{A}\subseteq\mathcal{S}} \frac{(|\mathcal{A}|-1)!(n-|\mathcal{A}|)!}{n!} \left(v(\mathcal{A}) - v(\mathcal{A}\backslash\{s_i\}\right),\ i = 1,\ldots,n. \tag{3}$$

Equation (3) shows how the Shapley value represents the average marginal individual contribution over all the possible different permutations in which the grand coalition $\mathcal{S}$ may be formed (Basili & Chateauneuf, 2020).[11]

Our problem consists of two states of the world ($n = 2$), this is, whether a tipping point occurs (Tip) or it does not occur ($No\ Tip$). The two states of the world occur with, say, probability Π_1^v and $\Pi_2^v = 1 - \Pi_1^v$, respectively. Let us now define the random variables $T \in \{Tip, No\ Tip\}$ denoting the occurrence (or not) of a tipping point and $C \in \{Low, Medium, High\}$ representing possible temperature scenarios as defined in Section 1. Hence, we have the following Bayes rule

$$P(T = Tip|C) = \frac{P(C|T = Tip)P(T = Tip)}{P(C|T = No\ Tip)P(T = No\ Tip) + P(C|T = Tip)P(T = Tip)}. \tag{4}$$

The Bayes rule in equation 4 can be used to update the probability of occurrence of a tipping point obtained via the Steiner point. Specifically, we choose Π_1^v for $P(T = Tip)$ as a prior probability, while $P(C|T = Tip)$ and $P(C|T = No\ Tip)$ can be obtained from expert knowledge. Alternatively, we can feed the formula a grid of values. In this last case the posterior probability $P(T = Tip|C)$ can be graphically represented as a surface. For the empirical analysis in Section 4 we opt for the latter.

[11] Recent research suggests that, for independent inputs, the Steiner point-Shapley value is bracketed between two different Sobol' indices. This result seems to hold also for the case of dependent inputs or expert judgments (Song, Nelson, & Staum, 2016, see also Owen and Prieur (2017) for further uses of the Shapley value in the context of ANOVA). Sobol' index provides a measure of the importance of inputs to a function and is defined in terms of the functional analysis variance decomposition (Sobol', 1990, 1993).

4 Eliciting Probabilities of Tipping Points by the Steiner Point

Kriegler et al. (2009) elicited beliefs of experts about the probability of triggering major changes in the Earth system associated to seven tipping points (Table 1 reports the tipping points under analysis) for three different global median temperature scenarios. As previously mentioned in Section 1, the three scenarios consider the crossing of a tipping point for a given temperature scenario and a fixed time horizon. Specifically, *Low* refers to an increase between about 1°C and 2°C by year 2200 in comparison with year 2000, *Medium* refers to an increase between about 2°C and 4°C by year 2200 in comparison with year 2000 and *High* refers to an increase between about 4°C and 8°C by year 2200 in comparison with year 2000. In light of the new evidence produced in over a decade, Armstrong McKay et al. (2022) classify new and known tipping elements as tipping points that act on a global scale or only locally (such as the dieback of Boreal forests, BOFO in Table 1). For other phenomena such as El Niño (NINO in Table 1), there seems to be insufficient evidence for it to be characterized by a tipping element. The decline of the ocean carbon sink (DOCS in Table 1), which was previously considered a tipping element, is categorized as threshold-free. For consistency of exposition with respect to Kriegler et al.'s data, we consider NINO and DOCS as if they were tipping elements.[12]

The raw data collected by Kriegler et al. (2009) are summarized in the dumbbell plots in Figures A1 to A7 in Appendix A. Each figure consists of three plots and each plot refers to one of the three temperature scenarios. To every row in the plots we attach one single expert. By direct inspection, we note, as also stressed in Kriegler et al. (2009), the tendency of the experts to place high probability in the high temperature scenarios.[13]

[12] Kriegler et al. (2009) report that BOFO and DOCS "were judged by experts to be of more speculative nature"

[13] The data are collected from Kriegler et al. (2009) supplemental appendix and are available at Federico Crudu's personal webpage (link) along with the replication code.

AMAZ	Dieback of the Amazon rainforest.
BOFO	Dieback of Boreal forests.
AMOC	Reorganization of the Atlantic meridional overturning circulation.
DAIS	Disintegration of the West Antarctic ice sheet.
MGIS	Melt of the Greenland ice sheet.
DOCS	Decline in ocean carbon sink.
NINO	Shift to a more persistent El Niño regime.

Table 1 Tipping points considered in Kriegler et al. (2009).

In order to provide policy relevant information Kriegler et al. (2009) combine, using an aggregation rule, the probability intervals supplied by the experts for every tipping point and every temperature scenario.[14] Their aggregation rule produces upper and lower probabilities. Our approach is substantially different. In fact, the Shapley value described in equation 3 returns the probability of occurrence of a tipping point conditional on a given temperature scenario.

Figure 1 features the probabilities associated to every tipping point and each climate scenario resulting from the aggregation of the experts' opinions via the application of equation 3. In Figure 1 we see that, in general, a high climate change scenario (red triangle) is more likely to trigger a tipping point. While this result is not too surprising, we also observe that there is a certain degree of heterogeneity across scenarios and tipping points. Specifically, AMOC has a very low chance to occur for the low climate change scenario. All the other tipping points have, for the same scenario a non negligible (say, larger than 10%) probability to occur. As we consider the higher climate change scenarios, $Medium$ and $High$, the probability of having a tipping point increases. It is remarkable that for $High$, the probabilities of occurrence for BOFO, DAIS, DOCS and MGIS are very high.

To provide further intuition for the interpretation of the plots in Figure 1, we adapt the terminology used in Kriegler et al. (2009). This is, we label as *remote* the prospect of having a tipping point if $\Pi_1^v < 0.1$, *significant* if $\Pi_1^v \geq 0.1$ and *large* if

[14] Kriegler et al. (2009) introduce a novel aggregation rule called forced consensus pooling. See also Clemen and Winkler (1999) and Nau (2002).

$\Pi_1^v \geq 0.5$. Under these criteria only AMOC for the low temperature scenario has a remote chance of being triggered. DAIS and MGIS have a large probability of being triggered under the high temperature scenario, while BOFO and DOCS have a large probability of being triggered under both the medium and high temperature scenario. The remaining scenarios have a significant probability of triggering a tipping point.

By applying the Bayes rule in equation 4 we update the results obtained via the Steiner point to generate tipping point scenarios for different combinations of $P(C|T = No\,Tip)$ and $P(C|T = Tip)$. Every point in the resulting surfaces (figures 2 to 8) is the conditional probability of crossing a tipping point given the likelihood of the (conditional) occurrence of the temperature scenario. We notice that the high temperature scenario produces large probabilities for the occurrence of tipping points. This result is particularly clear for BOFO (figure 3 (c)), AMOC (figure 4 (c)), DAIS (figure 5 (c)), DOCS (figure 6 (c)) and MGIS (figure 7 (c)). Committing to low climate change scenarios may produce remote probabilities of tipping, at least for AMOC (figure 4 (a)) and DAIS (figure 5 (a)), yet in most of the remaining cases the probabilities of triggering a tipping point are generally at least significant.

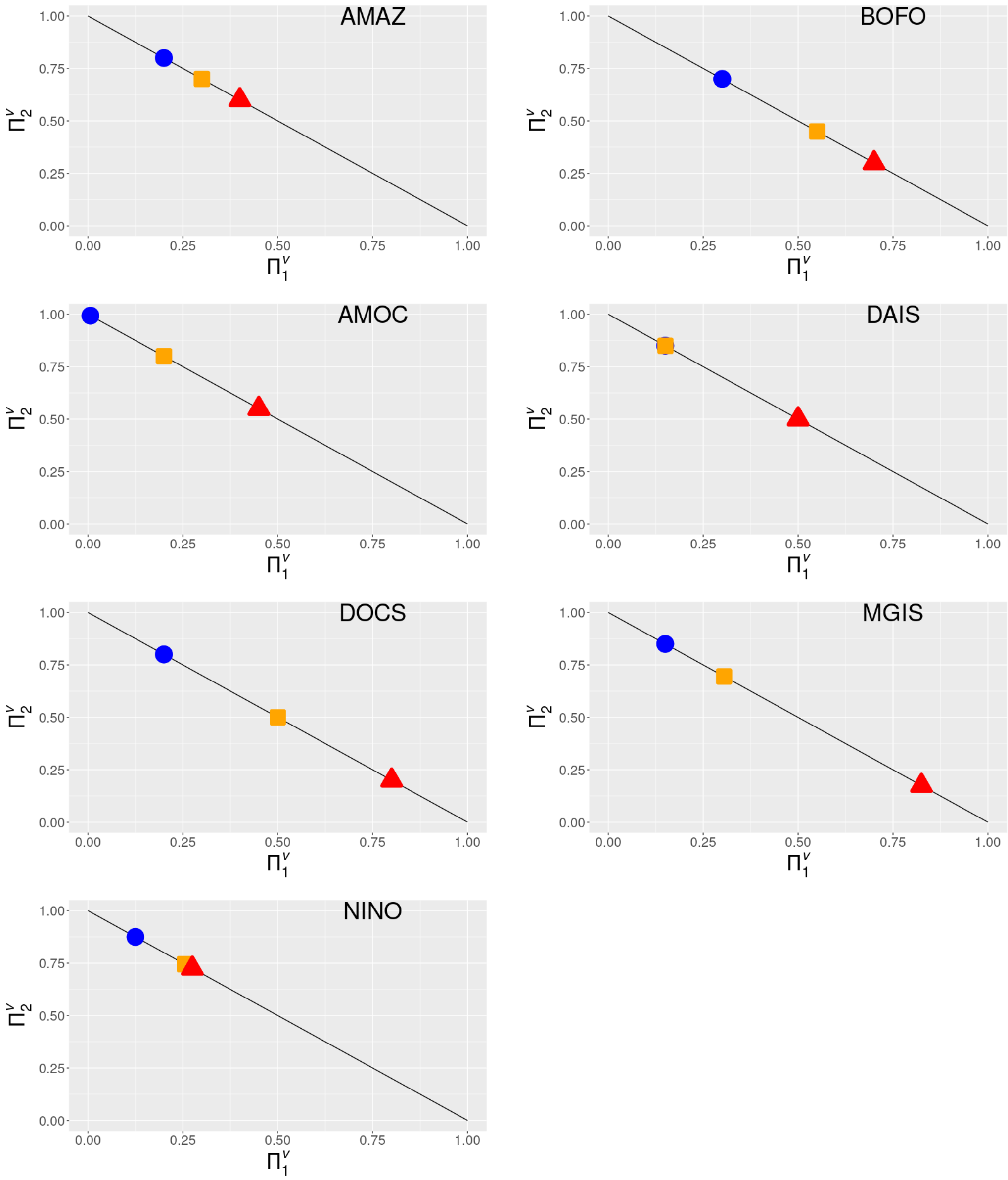
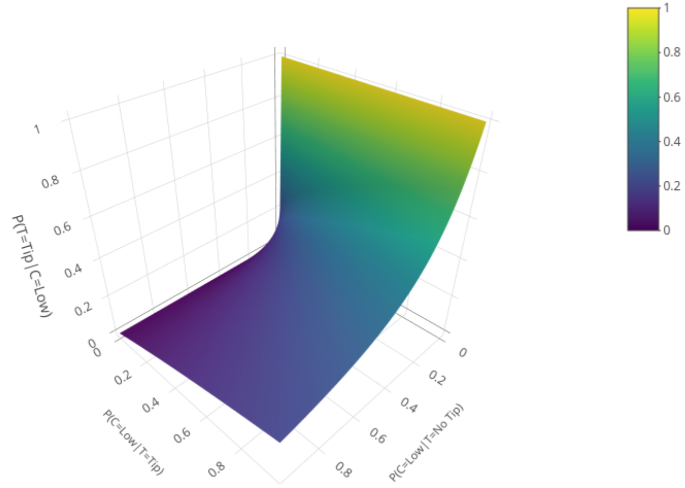
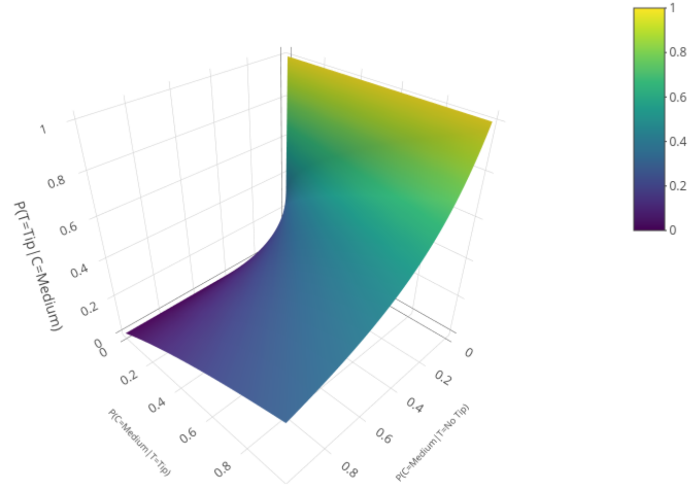


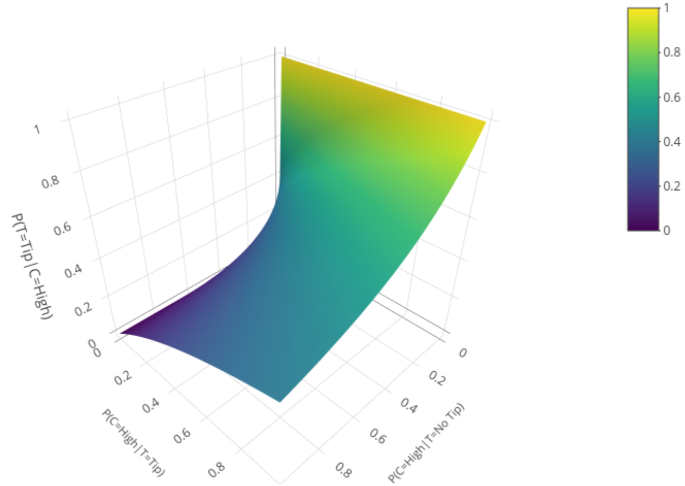
Fig. 1 Steiner points. The Low, Medium and High climate change scenarios are denoted by a circle, square and triangle respectively. Π_1^v is the probability that the tipping point is triggered, while Π_2^v is the probability that the tipping point is not triggered.



(a) Low temperature scenario.

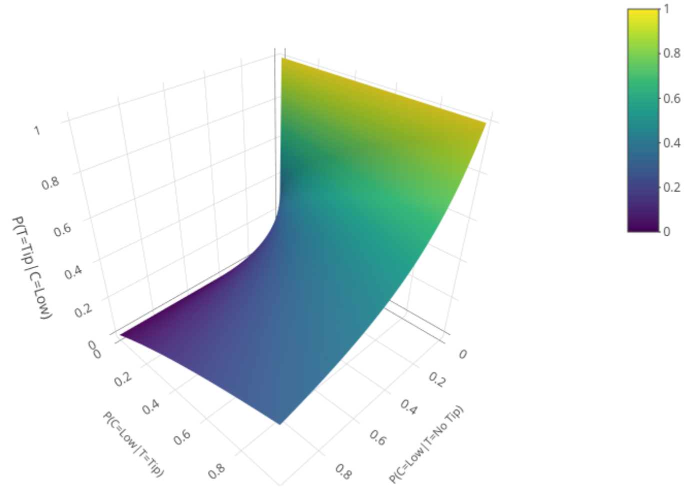


(b) Medium temperature scenario.

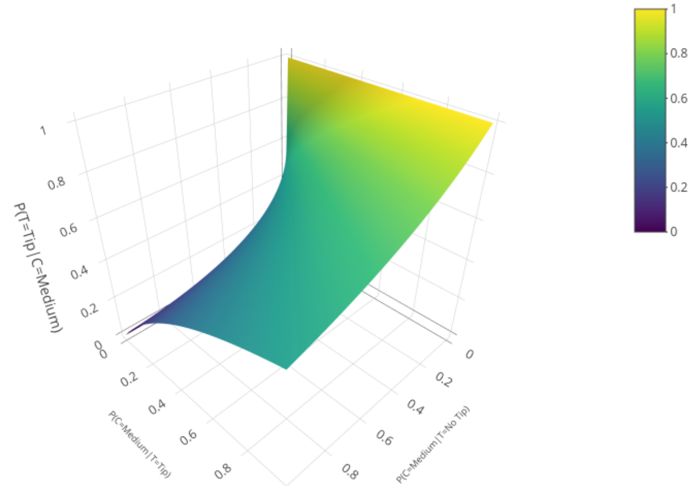


(c) High temperature scenario.

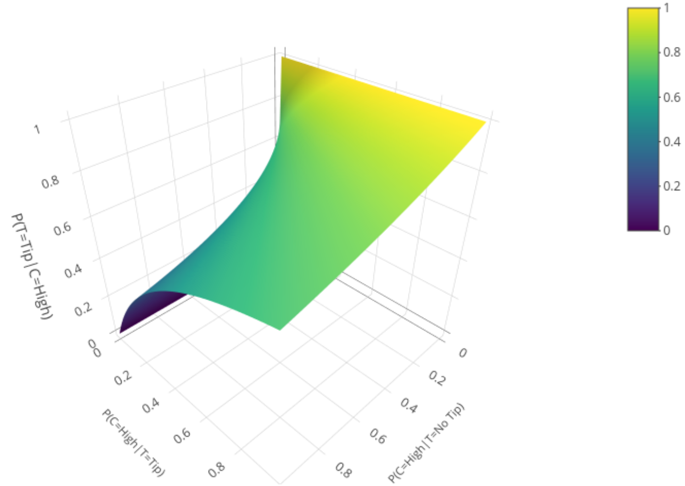
Fig. 2 Posterior surfaces for the AMAZ tipping point.



(a) Low temperature scenario.

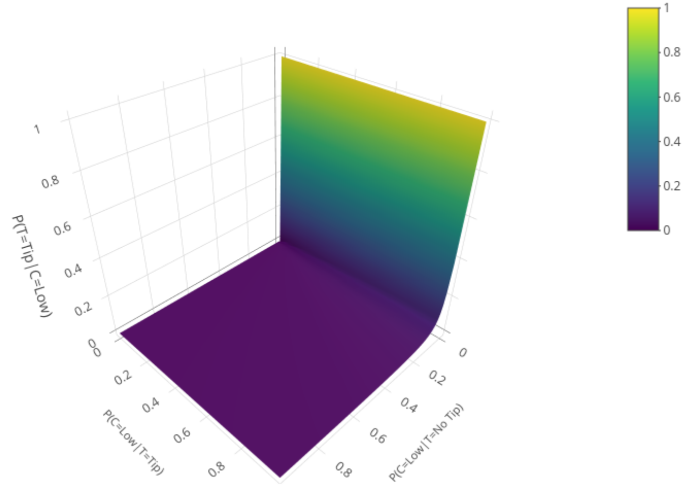


(b) Medium temperature scenario.

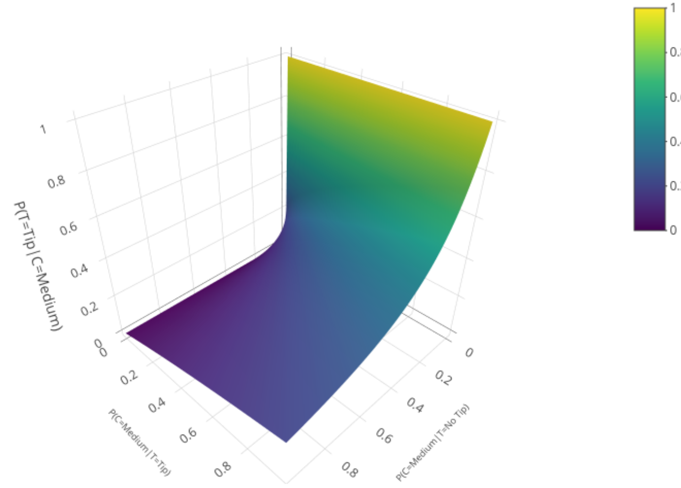


(c) High temperature scenario.

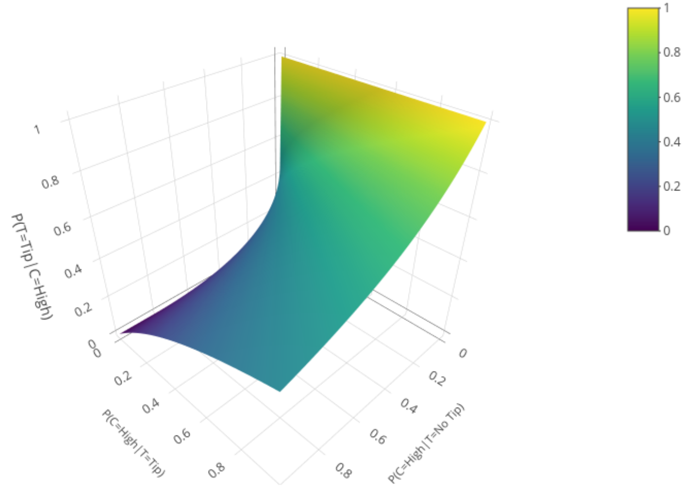
Fig. 3 Posterior surfaces for the BOFO tipping point.



(a) Low temperature scenario.

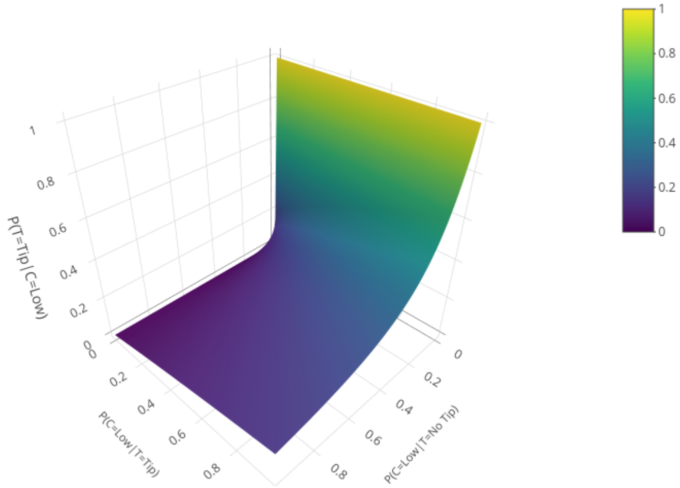


(b) Medium temperature scenario.

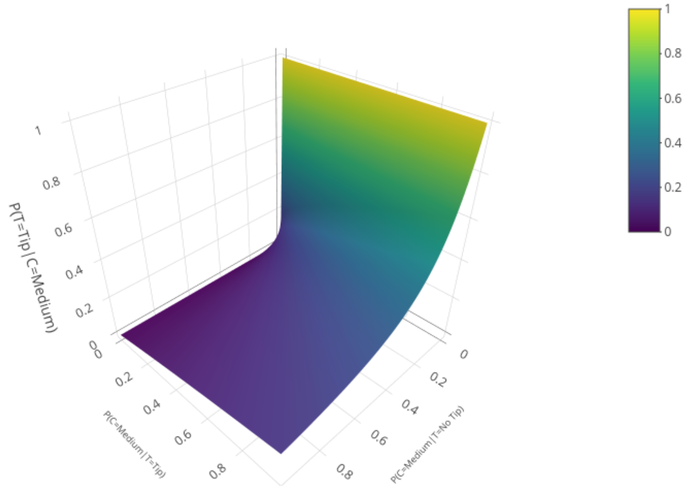


(c) High temperature scenario.

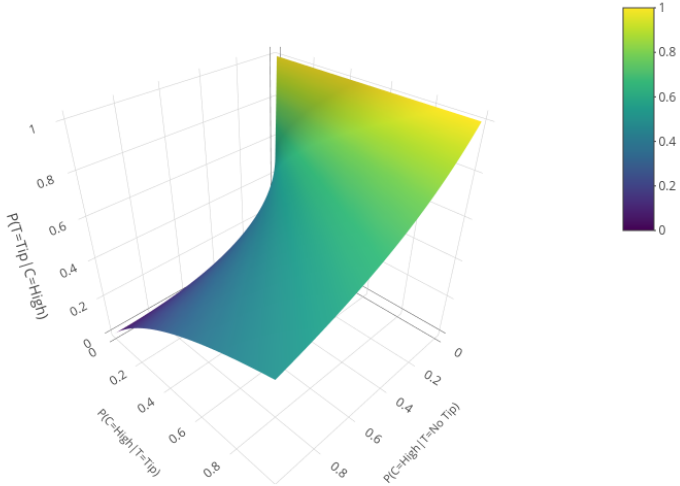
Fig. 4 Posterior surfaces for the AMOC tipping point.



(a) Low temperature scenario.

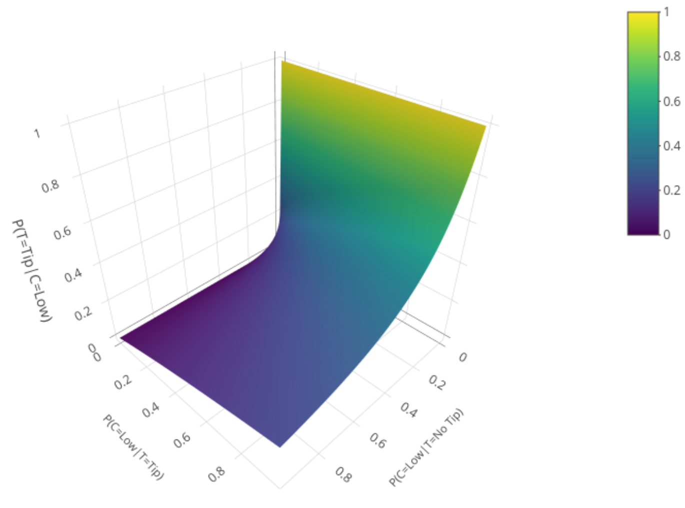


(b) Medium temperature scenario.

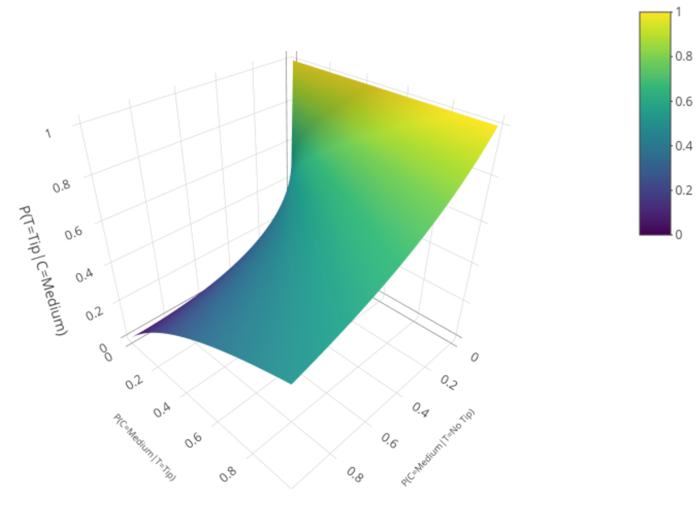


(c) High temperature scenario.

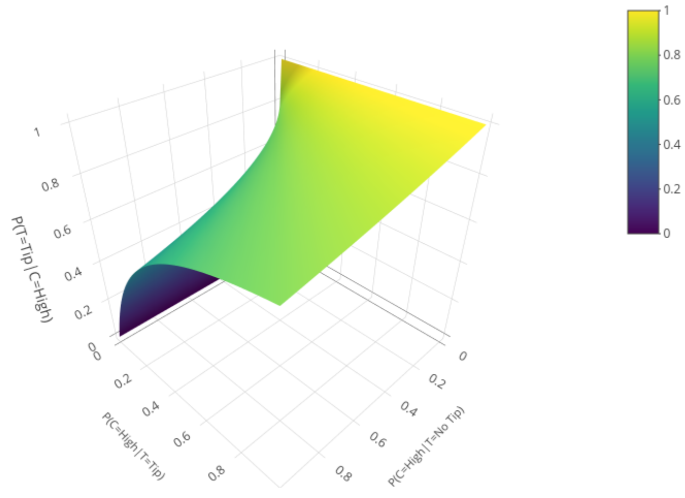
Fig. 5 Posterior surfaces for the DAIS tipping point.



(a) Low temperature scenario.

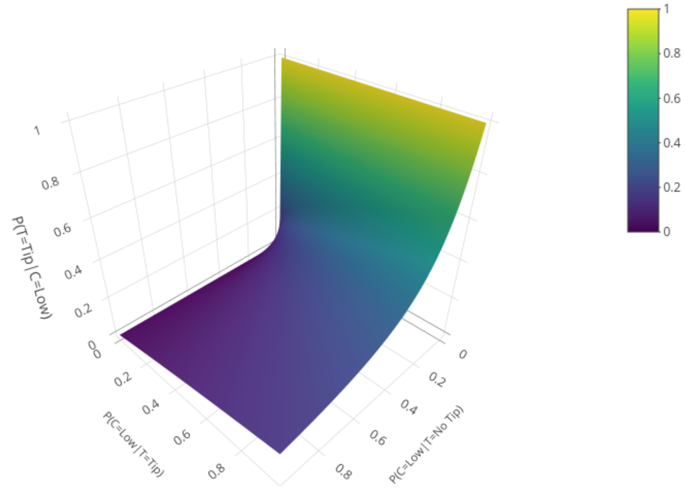


(b) Medium temperature scenario.

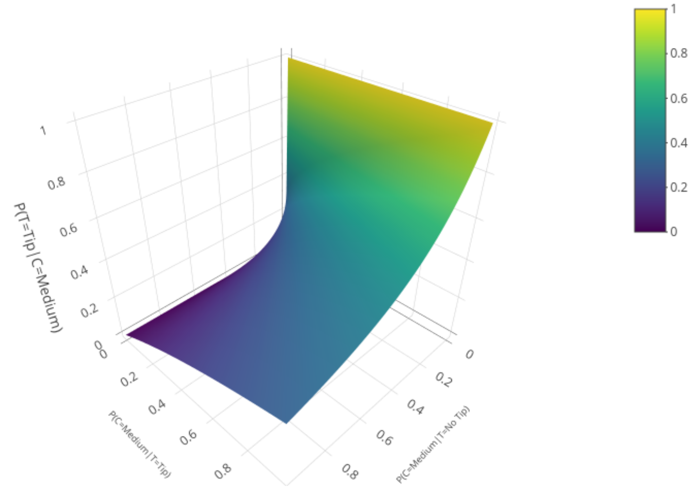


(c) High temperature scenario.

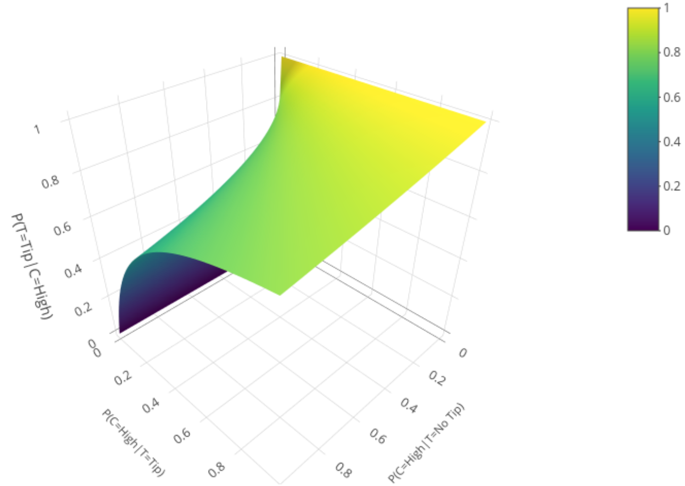
Fig. 6 Posterior surfaces for the DOCS tipping point



(a) Low temperature scenario.

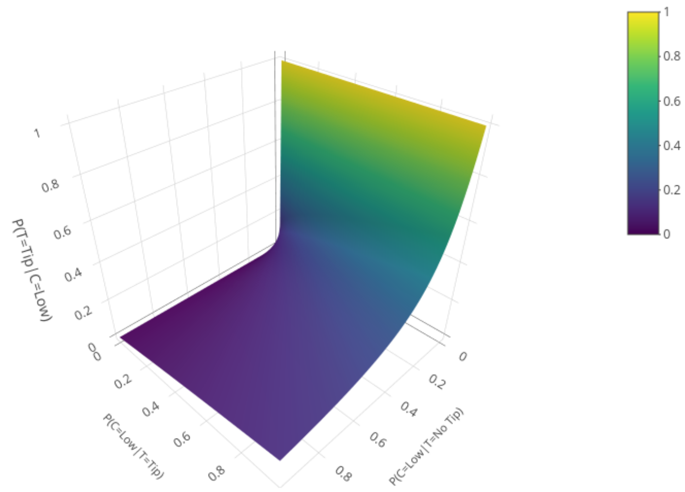


(b) Medium temperature scenario.

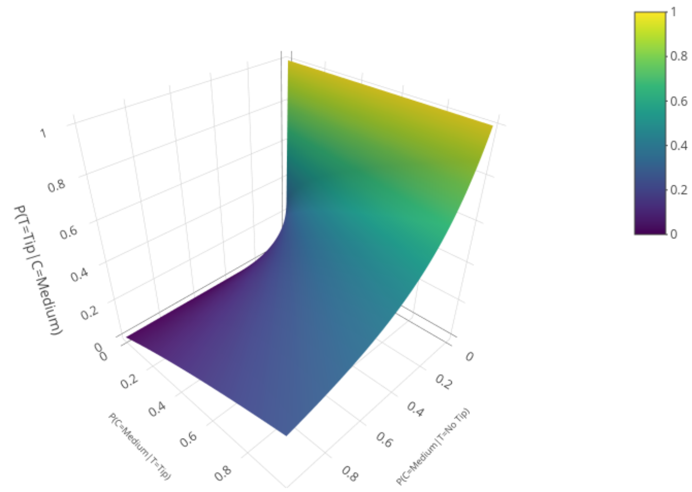


(c) High temperature scenario.

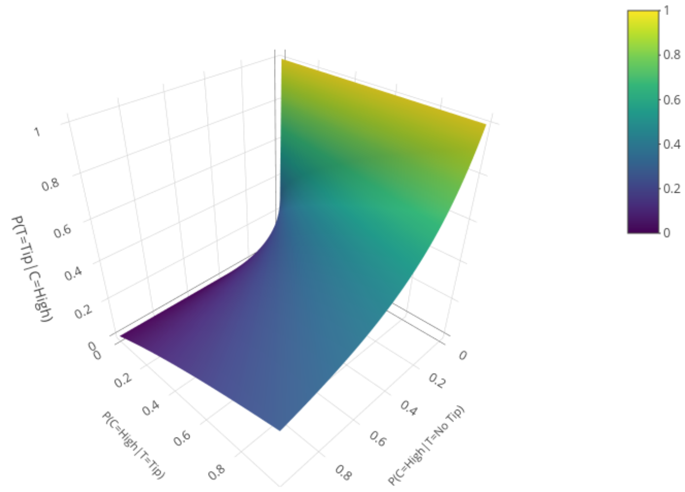
Fig. 7 Posterior surfaces for the MGIS tipping point.



(a) Low temperature scenario.



(b) Medium temperature scenario.



(c) High temperature scenario.

Fig. 8 Posterior surfaces for the NINO tipping point.

5 Conclusions

This paper considers the aggregation of experts' opinions about uncertain climate change scenarios. The aggregation rule based on the Steiner point provides an evaluation of the probability of occurrence of a tipping point.

In this context, the PM can evaluate the individual consistency (I-consistency) of an expert, that is, whether an expert is able to provide coherent probabilistic evaluations of possible future events, and the consistency of the whole group (G-consistency), that is, whether the intersection of the probability sets associated with the pool of experts is non empty. If the latter condition is verified the experts in the pool share common opinions about future states of the world. The Steiner point allows to automatically detect both individual and group consistency, but, more relevantly, it can be updated by the Bayes rule, whenever new information is available, without repeating a new trial of interviews among experts.

Ultimately, our results suggests that tipping points have higher probabilities to occur under high temperature scenarios. If taken in isolation this result is perhaps not surprising. A more interesting and concerning outcome is that the probability of crossing a tipping point threshold is noticeably different from zero, even under lower climate change scenarios and for the majority of tipping points under investigation. The probability surfaces obtained via the Bayes rule show that committing to a low climate change scenario may produce remote probabilities of realization of tipping points, yet in most cases such probabilities tend to be larger than 10%. Significantly, despite the fact that the data are over a decade old, to a large extent the results obtained with our methodology overlap with those produced in Armstrong McKay et al. (2022), even for the non occurrence of a tipping point for El Niño.

Due to limited data availability, we can only recover the marginal distribution of tipping points. We update the resulting probabilities using a standard Bayes rule for all the possible combinations of the likelihood of occurrence of a certain temperature

scenario. This approach generates posterior probability surfaces that may be interpreted as tipping points scenarios. The results suggest that some tipping points (the dieback of the Boreal forest, the reorganization of the Atlantic meridional overturning circulation, the decline of the ocean carbon sink and the melt of the West Antarctic and Greenland ice) are extremely likely to occur under the high temperature scenario. On the other hand, only committing to low climate change temperature corridors would substantially reduce the probability of occurrence of most tipping points.

Our analysis, consistently with the recent literature on the topic, highlights the fact that these type of events may turn out to be one of the main climatic emergencies in the upcoming years. Future research can further contribute to the understanding of these issues by addressing problems concerning elicitation data, the construction of joint distributions for tipping events and the estimation of causal relationships among tipping points.

References

Armstrong McKay, D.I., Staal, A., Abrams, J.F., Winkelmann, R., Sakschewski, B., Loriani, S., ... Lenton, T.M. (2022). Exceeding 1.5 c global warming could trigger multiple climate tipping points. *Science*, *377*(6611), eabn7950,

Basili, M., & Chateauneuf, A. (2011). Extreme events and entropy: a multiple quantile utility model. *International Journal of Approximate Reasoning*, *52*(8), 1095–1102,

Basili, M., & Chateauneuf, A. (2016). Aggregation of Coherent Experts' Opinions: A Tractable Extreme-Outcomes Consistent Rule. *Fusion methodologies in crisis management* (pp. 423–433). Springer.

Basili, M., & Chateauneuf, A. (2020). Aggregation of experts' opinions and conditional consensus opinion by the Steiner point. *International Journal of Approximate Reasoning*, *123*, 17–25,

Basili, M., & Pratelli, L. (2015). Aggregation of not independent experts': opinions under ambiguity. *Structural Safety*, *52*, 144–149,

Borsuk, M.E., & Tomassini, L. (2005). Uncertainty, imprecision, and the precautionary principle in climate change assessment. *Water Science and Technology*, *52*(6), 213–225,

Chateauneuf, A., & Cornet, B. (2018). Choquet representability of submodular functions. *Mathematical Programming*, *168*(1-2), 615–629,

Clemen, R.T., & Winkler, R.L. (1999). Combining probability distributions from experts in risk analysis. *Risk analysis*, *19*(2), 187–203,

Clemen, R.T., & Winkler, R.L. (2007). Aggregating probability distributions. W. Edwards, R.F. Miles, & D. von Winterfeldt (Eds.), *Advances in Decision Analysis: From Foundations to Applications.* Cambridge University Press.

Colson, A.R., & Cooke, R.M. (2018). Expert elicitation: using the classical model to validate experts' judgments. *Review of Environmental Economics and Policy*, *12*(1), 113–132,

Cooke, R. (2013). Uncertainty analysis comes to integrated assessment models for climate change…and conversely. *Climatic change*, *117*(3), 467–479,

Dansgaard, W., Johnsen, S.J., Clausen, H.B., Dahl-Jensen, D., Gundestrup, N., Hammer, C., … others (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, *364*(6434), 218–220,

De Campos, L.M., Huete, J.F., Moral, S. (1994). Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *2*(02), 167–196,

Frederikse, T., Buchanan, M.K., Lambert, E., Kopp, R.E., Oppenheimer, M., Rasmussen, D., van de Wal, R.S. (2020). Antarctic ice sheet and emission scenario controls on 21st-century extreme sea-level changes. *Nature communications*, *11*(1), 1–11,

Gaucherel, C., & Moron, V. (2017). Potential stabilizing points to mitigate tipping point interactions in earth's climate. *International Journal of Climatology*, *37*(1), 399–408,

Hall, J., Fu, G., Lawry, J. (2007). Imprecise probabilities of climate change: aggregation of fuzzy scenarios and model uncertainties. *Climatic Change*, *81*(3-4), 265–281,

Heinrich, H. (1988). Origin and consequences of cyclic ice rafting in the northeast atlantic ocean during the past 130,000 years. *Quaternary research*, *29*(2), 142–152,

IPCC (2018a). *Global Warming of 1.5° C. An IPCC Special Report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.* IPCC-World Meteorol. Organ Geneva.

IPCC (2018b). Summary for Policymakers. V. Masson-Delmotte et al. (Eds.), *Global Warming of 1.5° C. An IPCC Special Report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission*

pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. IPCC-World Meteorol. Organ Geneva.

Knutti, R., Allen, M.R., Friedlingstein, P., Gregory, J.M., Hegerl, G.C., Meehl, G.A., … others (2008). A review of uncertainties in global temperature projections over the twenty-first century. *Journal of Climate*, *21*(11), 2651–2663,

Kriegler, E., Hall, J.W., Held, H., Dawson, R., Schellnhuber, H.J. (2009). Imprecise probability assessment of tipping points in the climate system. *Proceedings of the national Academy of Sciences*, *106*(13), 5041–5046,

Kriegler, E., & Held, H. (2005). Utilizing belief functions for the estimation of future climate change. *International journal of approximate reasoning*, *39*(2-3), 185–209,

Lam, V., & Majszak, M.M. (2022). Climate tipping points and expert judgment. *Wiley Interdisciplinary Reviews: Climate Change*, *13*(6), e805,

Lenton, T.M., & Ciscar, J.-C. (2013). Integrating tipping points into climate impact assessments. *Climatic Change*, *117*(3), 585–597,

Lenton, T.M., Rockström, J., Gaffney, O., Rahmstorf, S., Richardson, K., Steffen, W., Schellnhuber, H.J. (2019). *Climate tipping points? Too risky to bet against.* Nature Publishing Group.

Marchau, V.A., Walker, W.E., Bloemen, P.J., Popper, S.W. (2019). *Decision making under deep uncertainty: from theory to practice*. Springer Nature.

Mastrandrea, M.D., Mach, K.J., Plattner, G.-K., Edenhofer, O., Stocker, T.F., Field, C.B., … Matschoss, P.R. (2011). The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups. *Climatic Change*, *108*(4), 675,

Nau, R.F. (2002). The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference*, *105*(1), 265–282,

Owen, A.B., & Prieur, C. (2017). On Shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, *5*(1), 986–1002,

Pechersky, S. (2015). A note on external angles of the core of convex tu games, marginal worth vectors and the weber set. *International Journal of Game Theory*, *44*(2), 487–498,

Rohmer, J., Le Cozannet, G., Manceau, J.-C. (2019). Addressing ambiguity in probabilistic assessments of future coastal flooding using possibility distributions. *Climatic Change*, *155*(1), 95–109,

Shapley, L.S. (1971). Cores of convex games. *International journal of game theory*, *1*(1), 11–26,

Sobol', I.M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, *2*(1), 112–118,

Sobol', I.M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment*, *1*(4), 407–414,

Song, E., Nelson, B.L., Staum, J. (2016). Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, *4*(1), 1060–1083,

Tomassini, L., Reichert, P., Knutti, R., Stocker, T.F., Borsuk, M.E. (2007). Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *Journal of Climate*, *20*(7), 1239–1254,

Wunderling, N., Donges, J.F., Kurths, J., Winkelmann, R. (2021). Interacting tipping elements increase risk of climate domino effects under global warming. *Earth System Dynamics*, *12*(2), 601–619,

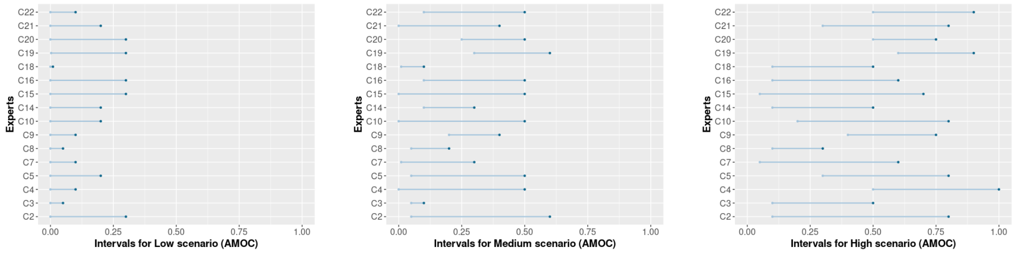
Zickfeld, K., Levermann, A., Morgan, M.G., Kuhlbrodt, T., Rahmstorf, S., Keith, D.W. (2007). Expert judgements on the response of the Atlantic meridional overturning circulation to climate change. *Climatic Change*, *82*(3-4), 235–265,

Zickfeld, K., Morgan, M.G., Frame, D.J., Keith, D.W. (2010). Expert judgments about transient climate response to alternative future trajectories of radiative forcing.

Proceedings of the National Academy of Sciences, *107*(28), 12451–12456,

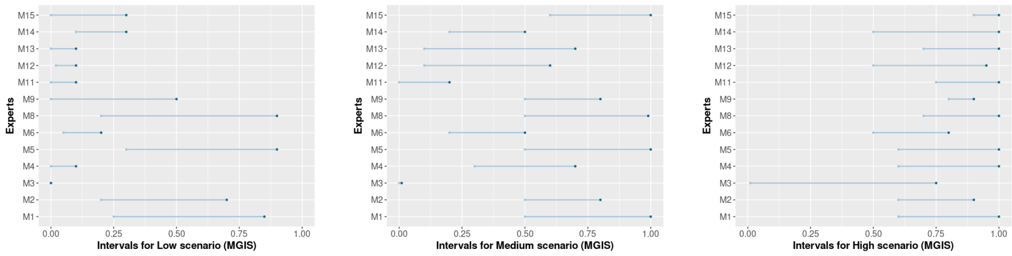
Appendix A Figures

This section displays a set of dumbbell plots that describe the data used in the analysis. The plots are similar to those in Figure 1 of Kriegler et al. (2009). Furthermore, in Kriegler et al. some of the experts are recognized as *core experts*. The caption of each plot indicates which experts are not core experts.



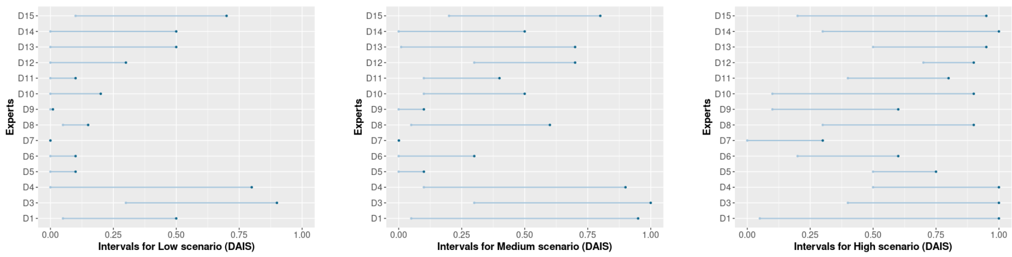
(a) Low (b) Medium (c) High

Fig. A1 Elicited probability intervals for the AMOC tipping point. In Kriegler et al. (2009) C9, C20 and C22 are not classified as core experts.



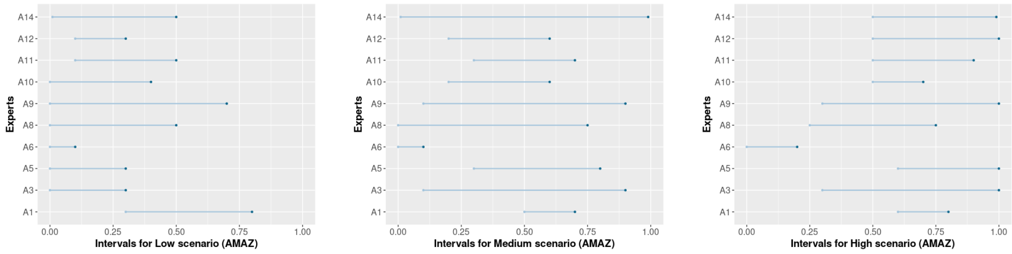
(a) Low (b) Medium (c) High

Fig. A2 Elicited probability intervals for the MGIS tipping point. In Kriegler et al. (2009) M1, M3 and M14 are not classified as core experts.



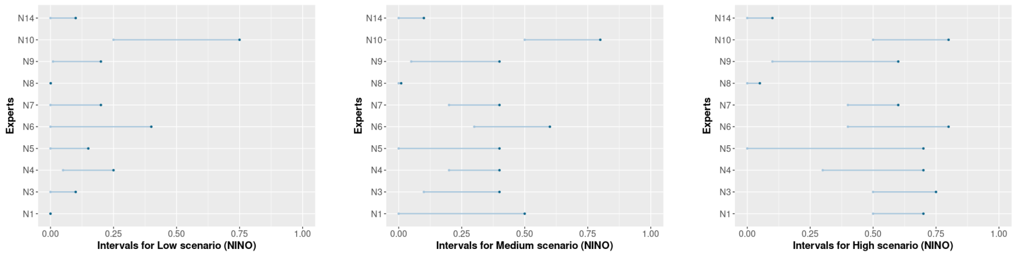
(a) Low (b) Medium (c) High

Fig. A3 Elicited probability intervals for the DAIS tipping point. In Kriegler et al. (2009) D7 and D10 are not classified as core experts.



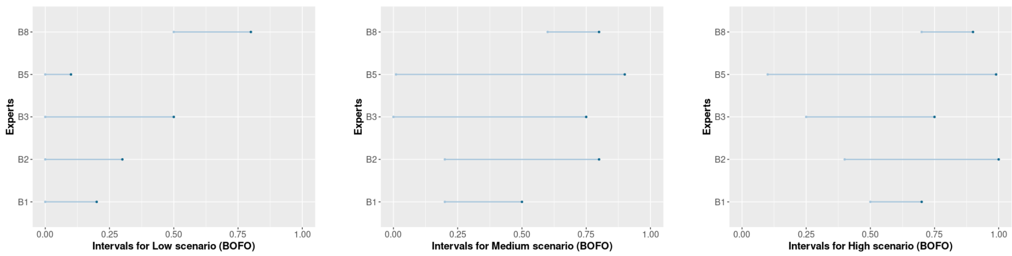
(a) Low (b) Medium (c) High

Fig. A4 Elicited probability intervals for the AMAZ tipping point. In Kriegler et al. (2009) A1, A6, A10 and A12 are not classified as core experts.



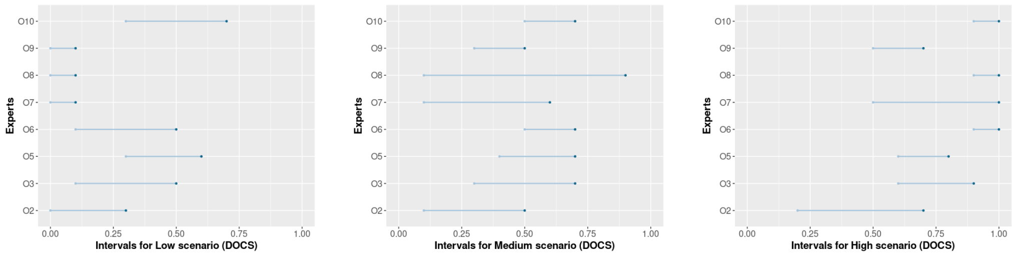
(a) Low (b) Medium (c) High

Fig. A5 Elicited probability intervals for the NINO tipping point. In Kriegler et al. (2009) N3, N6 and N10 are not classified as core experts.



(a) Low (b) Medium (c) High

Fig. A6 Elicited probability intervals for the BOFO tipping point. In Kriegler et al. (2009) B1 is not classified as a core experts.



(a) Low (b) Medium (c) High

Fig. A7 Elicited probability intervals for the DOCS tipping point. In Kriegler et al. (2009) O2 is not classified as a core experts.